

Mathematisches Argumentieren: Bedingungen und Wirkungen – eine Mixed Methods-Studie

Das vom Schweizerischen Nationalfonds unterstützte Projekt FEMAR (formatives Feedback zum mathematischen Argumentieren) untersuchte den Einfluss von Feedbackdialogen zwischen Lehrpersonen und Schüler*innen der 5. und 6. Klasse auf die schülerseitigen Kompetenzen beim mathematischen Argumentieren. Das Design enthielt eine 10-wöchige Intervention mit Aufgaben zum Argumentieren (Hess et al., 2020) und methodische Zugänge über Fragebögen für 71 Lehrpersonen und 1261 Schüler*innen, Leistungstests zum mathematischen Argumentieren vor und nach der Intervention, Interviews mit Lehrpersonen sowie Videoanalysen, u.a. mit einem literaturbasierten Rating der Feedbackdialoge.

Die folgende Mixed Methods-Studie (MMS; Schoonenboom & Johnson, 2017) beabsichtigt, markante quantitative Unterschiede qualitativ auszu-leuchten. Es werden Lehrpersonen-Interviews mit (i) auffällig tief und hoch eingeschätzten Dialogqualitäten und mit (ii) markant gesteigerten Klassenleistungen analysiert (vgl. Abb. 1). Wir stellen folgende Fragen: Wie erklären die Interviewaussagen unterschiedlich fremdeingeschätzte Dialogqualitäten? – Welche Haltungen und fachdidaktischen Überlegungen spiegeln sich in unterschiedlichen Qualitäten von Feedbackdialogen? – Wie erklären die Interviewaussagen auffällig starke Leistungssteigerungen?

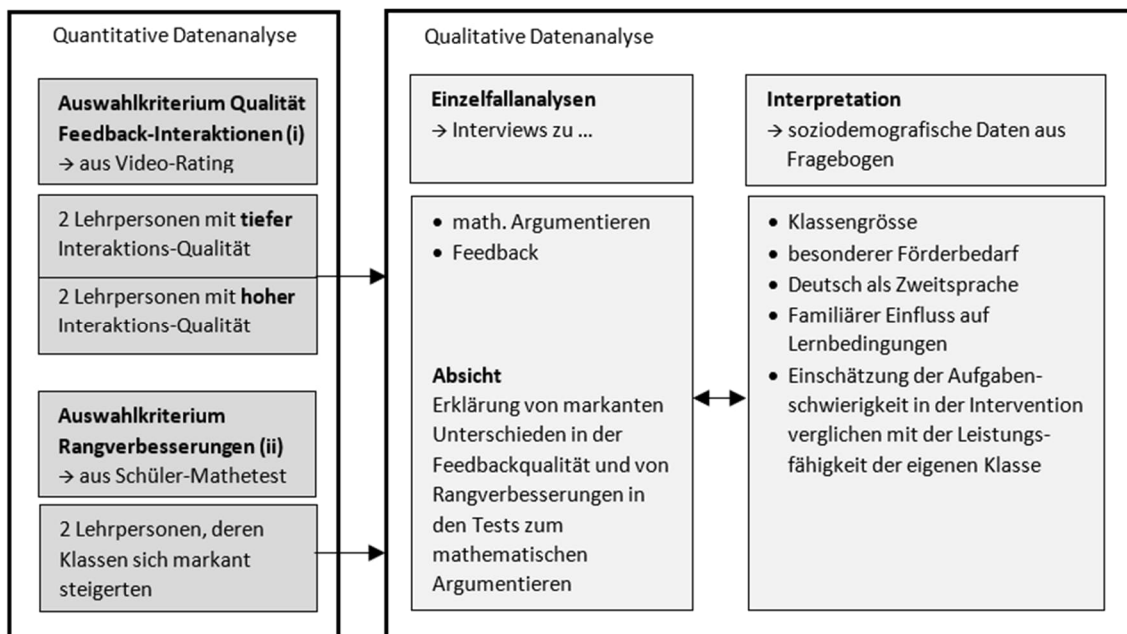


Abb. 1: Design der MMS zur Erklärung quantitativer Daten mit Interviewaussagen

Qualität von Feedback-Interaktionen und Tests zum Argumentieren

Lehr-/Lerndialoge tragen wesentlich zum Aufbau der mathematischen Argumentationsfähigkeit bei (Brunner, 2014; Howe & Abedin, 2013). Insofern sind Diskussionen, in welchen Lösungswege gesucht und Lösungen vermutet, erklärt und bewertet werden, typische Lernaktivitäten. Darin verifizieren Lernende mathematische Beziehungen und sie werden sich derer bewusst. Insofern ist das verbale Feedback von Lehrkräften ein Schlüssel zur Unterstützung des mathematischen Argumentierens (Fyfe & Brown, 2018).

In jeder Klasse ($n = 71$) wurde eine Lektion zum mathematischen Argumentieren videografiert, um die Qualität der Feedback-Interaktionen zu beurteilen. Die Einschätzung der Unterrichtsgespräche erfolgte über eine theorienbasierte Skala mit 13 Kategorien. Die Analyse der 44 bewerteten Videolektionen und der 13 Bewertungssitems erfolgte über eine Many-Facet-Rasch-Analyse (MRFM) mit Hilfe der FACETS-Software (Linacre, 2021). Das Rasch-Modell nutzt die Prinzipien der Intervallskalierung zur objektiven Messung von Daten, wobei die ordinalen Rohwerte in eine Reihe logarithmischer Transformationen einbezogen wurden, um Linearität zu erzeugen. Aus den Ergebnissen resultierte für jede Lehrperson ein Kompetenzwert für die Qualität der Feedbackgespräche und ein Schwierigkeitswert für jede Bewertungskategorie. Die Bewertungen unterscheiden sich auf einer Skala von 1 – 100 um 29 Punkte. Wir werteten die Interviews von zwei Lehrpersonen mit den tiefsten (30.28 und 31.99) und von zwei mit den höchsten Interaktionswerten (55.48, 58.40) qualitativ aus.

In einem Pre-Posttest-Design wurden die argumentativen Fähigkeiten zum mathematischen Argumentieren klassenbezogen ausgewertet. Der Vergleich zwischen den Testleistungen der Schüler*innen vor und nach der Intervention zeigt, dass die je beiden Klassen von Lehrpersonen mit sehr geringen und sehr hohen Interaktionsqualitäten vergleichbar zwischen 7 und 12 Ränge vorrückten. Die größten klassenbezogenen Leistungssteigerungen betragen beachtliche 25 und 31 Ränge. Sie führten zur Analyse von zwei weiteren Interviews (vgl. Abb. 1).

Diskussion der Ergebnisse

Wie erklären die Interviewaussagen unterschiedliche Dialogqualitäten?

Lehrpersonen mit *tiefer Dialogqualität* äußerten sich deutlich allgemeiner und abstrakter zum mathematischen Argumentieren und zu Feedback als diejenigen mit einer *hoch* eingestuften. Dennoch zeigen auch sie ein Bewusstsein und differenzierte Vorstellungen über wirksame Feedbacks. Im Vergleich zu den Lehrpersonen mit hoher Interaktionsqualität – welche fachdi-

daktisch plausible Beispiele anführten, von einer eigentlichen Feedbackkultur berichteten und sich am dialogischen Lernen orientieren – fallen ihre fachdidaktischen Unsicherheiten auf (z.B. keine bzw. falsche Unterscheidung zwischen Zahl und Ziffer oder unspezifische Verwendung von Darstellungen und Begriffen wie Argumentieren oder Darstellen). Die Interviews der Gruppe „Interaktionsqualität“ erklären weniger die Güte der Feedbackdialoge als die vergleichbare Verbesserung der klassenbezogenen Leistungen zum mathematischen Argumentieren.

Welche Haltungen und fachdidaktischen Überlegungen spiegeln sich in den unterschiedlichen Qualitäten von Feedbackdialogen?

In den Interviews fallen durchgehend differenzierte Absichten und Einstellungen auf. Es sind aber deutliche Unterschiede zwischen Lehrpersonen mit tief und hoch eingeschätzter Interaktionsqualität auszumachen. Letztere sprechen sozial-konstruktivistische bzw. kooperative Lernformen sowie eine überzeugende Feedbackkultur an. Deren Beispiele und die Werte der fremd eingeschätzten Dialogqualitäten lassen interpretieren, dass Gesagtes auch umgesetzt wird. Offenbar gelingt es ihnen, unterschiedlichen Lehr-/Lernbedingungen mit (fach-)didaktischem Geschick zu begegnen (Hess, 2003).

Wie erklären die Interviewaussagen auffällig starke Leistungssteigerungen?

Die beiden Lehrpersonen von Klassen mit *markanter Leistungssteigerung* im mathematischen Argumentieren überzeugen mit fachlicher Sicherheit und sozial-konstruktivistischen Argumenten wie dem „Austausch unter den Lernenden“, der „Veränderung der erklärenden Lehrerhaltung“, der „Lernbegleitung von Schüler*innen“ und „Feedbacks als Fragen formulieren“. Die Haltungen alleine vermögen die großen Leistungssteigerungen aber kaum zu erklären. Es braucht die Ergänzung mit den auffallend positiven Lehr-/Lernbedingungen: Die eine Lehrperson unterrichtet eine altersdurchmischte Klasse mit lediglich 10 Schüler*innen, kein Kind bedarf einer besonderen Förderung und lediglich zwei sind nicht deutscher Muttersprache. Sie stuft die familiäre Unterstützung als sehr gut ein. Es erstaunt also nicht, dass sie die eingesetzten Aufgaben als „zu einfach“ einschätzt und der Klasse eine gute Leistungsstärke zuspricht. Die Lehrerin stellt auch überzeugend dar, dass sie die günstigen Lehr-/Lernbedingungen – welche „eine intensivere Lernbegleitung ermöglichen“ – wirkungsvoll zu nutzen weiß. Die Lehr-/Lernbedingungen ihres Kollegen mit vergleichbar verbesserten Testleistungen weisen in dieselbe Richtung, obschon seine Klasse mit 26 Schüler*innen für Schweizer Verhältnisse eher groß ist.

Auch der Vergleich mit den Interviews von Lehrpersonen mit tiefen und hohen Dialogqualitäten führt nur zu einzelnen Begründungen, warum sich die

Klassen der Gruppe „Testleistungen“ dermaßen markant verbesserten. Es bedarf eines erneuten Blicks auf die Lehr-/Lernbedingungen, welche bei der Lehrerin mit geringer Interaktionsqualität bedeuten, dass sie eine Klasse mit 17 Schüler*innen unterrichtet, wovon sieben einer besonderen Förderung bedürfen und 10 nicht deutscher Muttersprache sind sowie ein belastendes familiäres Umfeld haben. Obschon sie die Testaufgaben als schwierig einstuft, verbesserten sich die Testleistungen ihrer Schüler*innen immerhin um 12 Ränge. Offenbar gelingt es ihr, auch unter ungünstigen Bedingungen wirkungsvolle Lernanlässe zu inszenieren. Die heterogenen Unterrichtsbedingungen der befragten Lehrpersonen plausibilisieren, dass aus den Interviews weder monokausale Hinweise auf unterschiedliche Interaktionsqualitäten noch auf markant verbesserte Testleistungen hervorgehen.

Die MMS vermag quantitative Daten mit fachdidaktischen Aspekten und unterrichtlichen Bedingungen zu ergänzen. Es liegt in der Differenziertheit qualitativer Analysen begründet, dass die Interpretationen eher zu feinen Grauabstufungen als zu Schwarz-Weiß-Kontrasten führen. Wir sind dennoch überzeugt, dass MMS zu valideren Ergebnissen und damit auch „zur Stärkung der Tiefe und Breite von Forschungsergebnissen“ beitragen (Buchholtz, 2021, S. 222).

Literatur

- Buchholtz, N. (2021). Voraussetzungen und Qualitätskriterien von Mixed-Methods-Studien in der mathematikdidaktischen Forschung. *Journal für Mathematikdidaktik*, 42(1), 219–245. <https://doi.org/10.1007/s13138-020-00173-0>
- Brunner, E. (2014). Verschiedene Beweistypen und ihre Umsetzung im Unterrichtsgespräch. *Journal für Mathematik-Didaktik*, 35(2), 229–249. <https://doi.org/10.1007/s13138-014-0065-6>
- Fyfe, E. R., & Brown, S. A. (2018). Feedback influences children's reasoning about math equivalence: A meta-analytic review. *Thinking & Reasoning*, 24(2), 157–178. <https://doi.org/10.1080/13546783.2017.1359208>
- Hess, K., Blum, V. & Smit, R. (2020). Argumentieren lernen mit Rubrics. Raster zur Steuerung und Beurteilung des mathematischen Argumentierens. *DiMawe – Die Materialwerkstatt*, 2(1), 49–62. <https://www.dimawe.de/index.php/dimawe/article/view/3590>
- Hess, K. (2003). *Lehren - zwischen Belehrung und Lernbegleitung. Einstellungen, Umsetzungen und Wirkungen im mathematischen Anfangsunterricht*. h.e.p.
- Howe, C., & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education*, 43(3), 325–356.
- Linacre, J. M. (2021). *Facets Rasch measurement computer program* (Version 3.83.6). <https://www.winsteps.com/facets.htm>
- Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69(2), 107–131. <https://doi.org/10.1007/s11577-017-0454-1>